A BAYENESIAN EXAMINATION OF AGRICULTURAL INVESTMENT

CHAD HART AND SERGIO H. LENCE

Proceedings of 46th Agricultural Finance Conference
"The Changing Nature of Agricultural Risks"
Delta Meadowvale Resort & Conference Centre
Mississauga, Ontario, Canada
October 4-6. 1999

University of Guelph

A BAYESIAN EXAMINATION OF AGRICULTURAL INVESTMENT

Chad Hart and Sergio H. Lence*

From a public policy standpoint, few issues are seen as crucial to enhance economic performance/growth as private business investment. In addition to its impact on long-term growth, private business investment contributes significantly to the short-term fluctuations in economic activity. It comes as no surprise that the theoretical and empirical analysis of private investment has occupied a prominent role in the research agenda of the economics profession for many years.

Following the thorough recent survey by Chirinko, investment models may be classified in two main categories: implicit and explicit. Among implicit models, the most important is the neoclassical model advanced by Jorgenson (1963, 1971). The two major explicit approaches are the q model advocated by Keynes, Tobin, and Brainard and Tobin, and the direct estimation of investment's Euler equation via generalized method of moments (GMM) (e.g., Hubbard and Kashyap). The theoretical paradigm underlying the neoclassical model is fundamentally different from the theory behind the q and the Euler equation-GMM models. In contrast, the q and the Euler equation-GMM models are but alternative empirical formulations of the same theory.

All three models have been used to study investment in agriculture. Articles relying on the neoclassical model include Weersink and Tauer, and Jensen, Lawson, and Langemeier. The q model was employed by Herendeen and Grisley, and Bierlen and Featherstone. Hubbard and Kashyap relied upon the Euler equation-GMM framework to analyze agricultural investment.

As noted by Chirinko, explicit models are theoretically more appealing than the neoclassical model. However, the latter typically performs better empirically. For this reason, alternative avenues have been explored to improve the empirical fit of explicit models. Quite possibly, the most successful of these alternatives has been the explicit consideration of financial market imperfections. The findings from this literature, reviewed recently by Hubbard, provide

*Post-Doctoral Research Associate and Associate Professor, respectively, Department of Economics, Iowa State University, Ames.

overwhelming indications that firms' internal financial variables have a significant impact on investment. Studies showing that agricultural firms' investment is affected by their financial situation include Jensen, Lawson, and Langemeier; Hubbard and Kashyap; Bierlen and Featherstone; Bierlen et al.; and Bierlen, Ahrendsen, and Dixon.

Upon review of the literature, a few stylized facts emerge. First, *there are no empirical studies testing whether the q model performs better than the neoclassical paradigm, or vice versa, in explaining agricultural investment*. This is surprising, given that knowing the relative empirical performance of the main alternative investment theories is highly relevant to analyze the likely impact of various policies. For example, policy-making inferences based on empirical estimates of the q paradigm might be seriously misleading if investment decisions are better represented by the neoclassical model.

Second, *the specific financial variables used to explain investment vary substantially across studies*. One might expect this because there is no theoretical justification to choose any particular financial variable over others. However, none of the studies applies formal model selection procedures to justify their financial variable choice. Unfortunately, one obvious implication is that the reported impact of financial variables on investment could simply be the consequence of data mining, and therefore lead to spurious conclusions.

Third, *none of the existing studies accounts for outlier effects*. This is unexpected, as it is well known that investment data exhibit an unusually large proportion of observations that might be defined as outliers. Some studies (e.g., Gilchrist and Himmelberg) follow ad hoc rules such as removing all of the observations below the 1st and above the 99th percentiles. However, it is unclear whether such ad hoc cutoff points are reasonable (e.g., the 5th and 95th quantiles, or even asymmetric cutoff points, might be warranted). Furthermore, the impact of such rules on statistical inference is unknown.

The present study contributes to the empirical investment literature by employing a Bayesian approach to address the aforementioned issues. More specifically, we (1) test the q model against the neoclassical model using data on investment in agricultural machinery and equipment for a panel of 366 Iowa farms over the period 1991-1998; (2) analyze the impact of financial variables on investment accounting specifically for model selection; and (3) incorporate outliers explicitly within the advocated modeling framework.

Importantly, the contribution of the present study transcends investment research. This is true because the techniques used originate from very recent works in Bayesian model selection, and our study is one of the first to employ them in an econometric setting. By showing how to integrate model selection and outlier detection procedures in a unified econometric framework, the advocated Bayesian approach should be of interest to applied researchers in many economic fields other than investment.

The results of our analysis provide more support for the neoclassical model than the q model. In fact, the q model is found to have very little support. Most of the support for the neoclassical model is embodied in the inclusion of the change in the cost of capital in the investment model. Financial variables, specifically lagged current asset values and off-farm income, also add greatly to the investment model. The addition of an outlier detection component to the model makes a significant difference in the results obtained and the inferences drawn.

Investment Models

In the interest of space, this section only provides a very brief sketch of the two competing models used for empirical analysis --neoclassical and q. Both models are developed in great depth in many excellent sources (e.g., Chirinko and references therein). In addition, financial and demographic variables used in econometric models of investment are briefly discussed.

The Neoclassical Model of Investment

In the neoclassical model, the firm is assumed to maximize discounted expected profits over an infinite horizon. Capital depreciates at a geometric rate, and there are no adjustment costs or vintage effects. The optimal physical capital level is determined by output and the user cost (rental price) of capital.

Assuming the production function exhibits a constant elasticity of substitution (σ) between variable inputs and capital, the investment equation for the neoclassical model may be written as follows:[1]

$$(1.1) \quad I_{i,t}/K_{i,t-1} = \alpha_0^{nc} + \alpha_1^{nc}\sum_{j=0}^{J}\gamma_j^{nc}\left(\frac{Q_{i,t-j}}{K_{i,t-1}C_{i,t-j}^{\sigma}} - \frac{Q_{i,t-j-1}}{K_{i,t-1}C_{i,t-j-1}^{\sigma}}\right),$$

where $I_{i,t} \equiv K_{i,t} - K_{i,t-1}$ denotes firm i's total net investment in period t, $K_{i,t}$ is the capital stock at the end of period t, α_0^{nc} is the depreciation rate, $Q_{i,t}$ represents output, $C_{i,t}$ is the user cost of capital divided by the output price, α_1^{nc} is a parameter from the production function, and γ_j^{nc} ($\Sigma\gamma_j^{nc} = 1$) are parameters representing capital delivery lags. According to the neoclassical model, total investment should be positively related to the initial capital stock and to the change in output (normalized by the initial capital stock), and negatively related to the change in the user cost of capital.

The q Model of Investment

The q model assumes that there are adjustment costs associated with the addition of new capital to the firm. In the standard case of quadratic adjustment costs, the corresponding investment equation is (1.2):

$$(1.2) \quad I_{i,t}/K_{i,t-1} = \alpha^q q_{i,t},$$

where parameter α^q denotes the inverse of the adjustment cost of new capital, and $q_{i,t}$ is the expected present value of the marginal product of new capital for firm i.

The $q_{i,t}$ term in (1.2) embodies all of the information available about the present value of adding one more unit of capital to the firm. For a given level of adjustment costs ($1/\alpha^q$),

[1] Superscripts are used to economize in symbols. In what follows, superscripts "*nc*" and "*q*" refer to the neoclassical and the *q* models, respectively.

investment (normalized by existing capital) increases with $q_{i,t}$. Greater adjustment costs (i.e., smaller α^q) imply a smaller response of (normalized) investment to $q_{i,t}$.

Financial and Demographic Variables

In models (1.1) and (1.2), internal and external sources of finance are treated as perfect substitutes. The firm is unconcerned or unaffected by the choice of internal or external funds. This would be true if there were no transaction costs or asymmetric information problems between lenders and borrowers. Although this type of assumption may be adequate in some settings, it is hard to justify for agricultural investment at the farm level. Numerous studies of agricultural investment provide evidence of financial constraints (e.g., Jensen, Lawson, and Langemeier; Hubbard and Kashyap; Bierlen and Featherstone; Bierlen et al.; Bierlen, Ahrendsen, and Dixon; and Benjamin and Phimister).[2]

Following the existing literature, it is hypothesized that a firm's financial situation does affect its investment behavior. Unfortunately, there is no single measure of a firm's financial situation. There are many financial indicators that provide partial information about a firm's financial situation. Theory provides no guidance as to which of such indicator(s) is/are the most appropriate to use, or how it/they should enter the investment equation. Examination of the previous investment literature reveals little consensus regarding the choice of financial indicators to include in the investment equation. Similar conclusions can be drawn from the agricultural credit literature (e.g., Miller and LaDue, and Knopf and Schoney), which is concerned with the most influential of the farm's financial variables for the credit decision from a lender's point of view.

Arbitrariness in the choice of financial indicators to include in the investment equation proves unavoidable. Further, special care has to be taken to minimize the problem of multicollinearity, which is a distinct characteristic of financial indicators. For these reasons, informed judgement was exercised to select the set of financial variables to be included in the

[2]However, Weersink and Tauer found that investment *decreases* with the level of real net farm income.

empirical specification. The selected set of financial indicators comprises the initial value of short-term assets (CA_{t-1}), initial net worth (NW_{t-1}), initial current liabilities (CL_{t-1}), initial total liabilities (TL_{t-1}), lagged off-farm income (OFI_{t-1}), and lagged farm net cash flow (NCF_{t-1}). If the firm's financial situation affects its investment, the latter is expected to be positively (negatively) impacted by *CA*, *NW*, *OFI*, and *NCF* (*CL* and *TL*).

Given that the firms analyzed here are proprietorships, socioeconomic characteristics are hypothesized to be important determinants of investment behavior as well. In particular, lifecycle considerations provide a theoretical justification for including farmer's age (*AGE*) as an explanatory variable in the investment regression. Bierlen and Featherstone; Bierlen, Ahrendsen, and Dixon; and Jensen, Lawson, and Langemeier found that a farmer's age significantly affected investment, whereas Weersink and Tauer and Bierlen et al. found no significant effects of age on investment. Other demographic characteristics (e.g., education and household size) are also likely to affect investment, but data limitations prevented us from considering them explicitly.

Data

The data employed originate from the Individual Farm Analysis data set of the Iowa Farm Business Association, for the years 1991 through 1998 (these were the years for which access to the individual records was allowed). The data are collected by Iowa Farm Business Association consultants, and kept on an inventory basis under standardized accounting procedures. For each year, the data set contains records on more than 700 variables, including detailed production and financial information, for over 1,000 Iowa farms. After combining the 1991 through 1998 data sets and removing the farms that did not have complete records for the whole period, 366 farms were left for the present analysis.[3]

Table 1 reports summary statistics for the study variables. After allowing for leads and lags, there are 2,196 observations left for each variable. The physical quantity of capital stock

[3]Upon consultation with the supervisor of the Iowa Farm Business Association data, 7 farms were also removed from consideration due to extremely large changes between previous end-of-year values and beginning-of-year values or to having investment ratios (the ratio of investment to the capital stock) greater than 5.

($K_{i,t}$) is obtained by dividing the value of machinery and equipment by the series "Producer Price Index for Agricultural Machinery" reported by the Bureau of Labor Statistics. Investment ($I_{i,t}$) is measured as the difference between the value of machinery and equipment at the beginning and the end of the year ($K_{i,t} - K_{i,t-1}$). Output ($O_{i,t}$) is divided into two measures, one for crop production ($CO_{i,t}$) and one for livestock production ($LO_{i,t}$). Both measures are calculated as the sum of the quantities of commodities produced on the farm multiplied by the average deflated price corresponding to those commodities over the study period (see Appendix A for details).

All monetary values, except where noted, are deflated using the series "Producer Price Index for Finished Goods" reported by the Bureau of Labor Statistics. User cost of capital (relative to output price) ($C_{i,t}$) is an index representing the price at which capital may be obtained.

Table 1. Summary statistics (all financial variables in 1998 dollars).[a]

Variable	Mean	Median	Standard Deviation	Minimum	Maximum
Investment ($I_{i,t}$)	5,862	−138	29,668	−240,365	293,239
Lagged machinery value ($K_{i,t-1}$)	128,367	104,675	90,160	1,372	677,348
Lagged investment ($I_{i,t-1}$)	5,951	−32	28,438	−240,365	293,239
Tobin's q (q)[b]	0.17	0.14	0.506	−1.19	7.61
Change in crop output ($\Delta CQ_{i,t}$)	6,232	4,265	62,792	−387,906	386,921
Change in livestock output ($\Delta LQ_{i,t}$)	1,565	0	54,854	−725,708	603,713
Change in cost of capital ($\Delta C_{i,t}$)	0.93	0.91	1.91	−5.21	7.47
Lagged operator age ($AGE_{i,t-1}$)	47.7	47	10.3	24	78
Lagged current assets ($CA_{i,t-1}$)	212,814	173,437	163,382	0	1,713,541
Lagged current liabilities ($CL_{i,t-1}$)	73,309	38,165	102,083	0	833,026
Lagged net worth ($NW_{i,t-1}$)	599,033	463,513	486,999	−68,437	3,531,546
Lagged total liabilities ($TL_{i,t-1}$)	193,566	143,979	198,524	0	1,330,822
Lagged off-farm income ($OFI_{i,t-1}$)	8,286	1,014	14,257	−45,216	142,712
Lagged net cash flow ($NCF_{i,t-1}$)	69,479	56,203	63,748	−550,218	612,725

[a]All of the neoclassical and financial variables are normalized by the value of lagged machinery during the estimation of the regression models.
[b]This variable is actually a multiple of Tobin's q. The constant depends on parameters from the production function.

Following Weersink and Tauer, it is computed from (2.1):

$$(2.1)\quad C_{i,t} = \frac{1}{p_t^O}\left(\frac{p_t^K}{1-m_{i,t}}\right)\left[(1-m_{i,t})\delta + r_t - \left(\frac{p_t^K - p_{t-1}^K}{p_t^K}\right)\right],$$

where p^O represents the output price index, p^K denotes the price of new capital at time t, $m_{i,t}$ is farmer i's marginal tax rate, δ is the capital depreciation rate, and r_t denotes the interest rate. The price of new capital is proxied by the "Producer Price Index for Agricultural Machinery." The depreciation rate is set at 15 percent per year. The marginal tax rate is based on federal and state tax laws and the farmer's net farm income for that year. The interest rate is the "Average Effective Interest Rate on Loans Made -- Farm Machinery and Equipment" reported in the *Agricultural Finance Databook*. Details about the construction of the output price index (p^O) are provided in Appendix A.

In (1.2), $q_{i,t}$ is unobservable. Following Abel and Blanchard, Gilchrist and Himmelberg, and Beirlen and Featherstone, $q_{i,t}$ is estimated by fitting the vector autoregression (VAR) (2.2),

$$(2.2)\quad Z_{i,t} = \Gamma Z_{i,t-1} + u_{i,t},$$

and using the estimated matrix of coefficients Γ to calculate $q_{i,t}$:

$$(2.3)\quad q_{i,t} = c_1\left[\underline{1} - \left(\frac{1-\delta}{1+r_t}\right)\Gamma\right]^{-1}\left(\frac{1-\delta}{1+r_t}\right)\Gamma^2 Z_{i,t-1}.$$

In (2.2), $u_{i,t}$ is an error term and $Z_{i,t}$ is a vector of current and lagged fundamentals that contains the marginal value product of machinery (in its first row) and other variables that help predict the marginal value product of machinery. In (2.3), $c_1 \equiv [1, 0, \ldots, 0]$ and $\underline{1}$ is a conformable identity matrix. The marginal value product of machinery is set equal to the average value product of machinery (which is valid if the production function is homogeneous of degree 1). The marginal value product of machinery is measured as the ratio of the sum of management returns and machinery depreciation to the value of the capital stock.

Besides the marginal value product of machinery, $Z_{i,t}$ includes all of the financial variables previously discussed (*CA*, *NW*, *OFI*, *NCF*, *TL*, and *CL*), to make sure that their effect in the q model is due to financial constraints and not to the fact that they help predict the marginal product of capital. The one other variable included in $Z_{i,t}$ is the ratio of the total value of production to the value of the capital stock. This ratio has been employed as an alternative measure of the marginal value product of capital (e.g., Gilchrist and Himmelberg).

Farmers in the present data set are a self-selected sample, as they have chosen to submit information to the Iowa Farm Business Association. To see how this self-selection might impact the analysis, the farms under study were compared with the set of all Iowa farms in the 1992 Census of Agriculture. The present data set overrepresents "typical" farms (those between 180 acres and 2000 acres), and underrepresents both small (less than 180 acres) and very large (more than 2000 acres) farms. Similar patterns emerge in the comparisons for machinery value and operator age.

Methods

The empirical analysis is based on a Bayesian approach to estimate the following linear (in the coefficients) regression:

(3.1) $I_{i,t}/K_{i,t-1} = \beta X_{i,t} + y_t + \varepsilon_{i,t}$,

where β is the vector of unknown coefficients to be estimated, $X_{i,t}$ is a vector of explanatory variables, y_t are annual effects ($y_t \sim$ iid N(0, σ_y^2) for all t), and $\varepsilon_{i,t}$ is an error term. Vector $X_{i,t}$ consists of a constant, explanatory variables in the neoclassical and q models, financial variables, and the demographic variable. To allow for nonlinear effects in the explanatory variables, cross-product and quadratic terms are also included in $X_{i,t}$. Vector $X_{i,t}$ includes lagged values of the dependent variable as well, yielding a vector of 73 terms (see Tables 2 or 4 for a complete list).[4]

[4]Note that cross product terms between explanatory variables for the neoclassical model (1.1) and q model (1.2) are omitted, because it is assumed that the two theoretical models are mutually exclusive. Also, the quadratic terms for q and outputs are omitted for consistency with theoretical models (1.2) and (1.1), respectively.

Regression model (3.1) is formulated as a hierarchical normal linear model. Since exploratory examinations indicate that outliers are likely to be present, residuals $\varepsilon_{i,t}$ are modeled as coming from a contaminated normal distribution:

$$(3.2) \quad \varepsilon_{i,t} \sim \begin{Bmatrix} N(0,\sigma_\varepsilon^{\ 2}) & \text{with probability } \eta \\ N(0,\kappa^2\sigma_\varepsilon^{\ 2}) & \text{with probability } (1-\eta) \end{Bmatrix},$$

where $\kappa^2 > 1$ is a variance-inflation parameter. This implies that the standard deviation of outliers is $|\kappa|$ times greater than the standard deviation of non-outliers.

A fundamental advantage of the Bayesian approach is that it provides a consistent framework for model selection. Here, this advantage we exploited to choose regressors by employing Geweke's variable selection method. Under this technique, a prior probability $\underline{p}_k$ is assigned to the event $\beta_k = 0$. Further, if $\beta_k \neq 0$, the prior distribution on β_k is assumed to be normal, $p(\beta_k) \propto \exp[-0.5\ (\beta_k/\tau_k)^2]$.

The distributional assumptions, combined with the regression equation (3.1), imply that $I_{i,t}/K_{i,t-1}|\, y_t, \beta, X_{i,t}, \sigma_\varepsilon^2, \sigma_y^2, \theta_{i,t}, \eta, \kappa \sim \mathrm{N}[y_t + X_{i,t}\beta,\ \sigma_\varepsilon^2(\theta_{i,t} + \kappa^2 - \kappa^2\theta_{i,t})]$, where $\theta_{i,t}$ is an outlier indicator defined as $\theta_{i,t} = 1$ if $\varepsilon_{i,t} \sim \mathrm{N}(0,\ \sigma_\varepsilon^2)$ and $\theta_{i,t} = 0$ otherwise. In the Bayesian framework, y_t, β, and σ_ε^2 are the model parameters, and σ_y^2, $\theta_{i,t}$, η, and κ are hyperparameters.

The joint posterior distribution of all parameters is obtained by combining the likelihood function with prior distributions for the parameters and hyperparameters. Prior distributions are as follows: $p(\sigma_\varepsilon^2) \propto 1/\sigma_\varepsilon^2$ for σ_ε^2, $\sigma_y^2 \sim \text{Inverse-}\chi^2(n_0, \sigma_0^2)$ for σ_y^2, $p(\theta_{i,t} \mid \eta) \propto \eta^{\theta_{i,t}}(1-\eta)^{(1-\theta_{i,t})}$ for $\theta_{i,t}$, and $p(\eta \mid \gamma, \varphi) \propto \eta^{(\gamma-1)}(1-\eta)^{(\varphi-1)}$ for η. That is, the prior distribution for σ_ε^2 is noninformative. The prior for σ_y^2 can be interpreted as adding n_0 observations with an average squared deviation of σ_0^2 to the analysis of σ_y^2. The assumed prior Binomial and Beta distributions for $\theta_{i,t}$ and η, respectively, allow us to set the values of (hyper) hyperparameters γ and φ so as to reflect our prior beliefs about the proportion of potential outliers in the data set.

To assess the sensitivity of the results to the priors for the outlier detection and variable selection components, estimation is performed under ten combinations of priors. For the outlier detection component, three sets of priors are used; namely, that 10% ($\gamma = 18$, $\varphi = 2$), 50% ($\gamma = 1$, $\varphi = 1$), and 90% ($\gamma = 2$, $\varphi = 18$) of the observations are outliers. For the variable selection component, we also employ three priors: 10% ($\underline{p}_k = 0.9$), 50% ($\underline{p}_k = 0.5$), and 90% ($\underline{p}_k = 0.1$) probability that each regressor (main and cross effects) is included in the model. All nine combinations of these priors are examined, along with a run assuming that the data set contains no outliers and that each variable has a 90% ($\underline{p}_k = 0.1$) prior probability of being included in the model. In total, 10 separate estimation runs are performed.

The prior $\sigma_y^2 \sim \text{Inverse-}\chi^2(20, 0.01)$ is chosen for σ_y^2. This highly informative prior is based on estimates from a classical analysis of the model, and effectively adds 20 observations with an average squared deviation of 0.01 to the analysis of σ_y^2. A highly informative prior for σ_y^2 is used to alleviate the problem of separating the intercept from the random annual effects, due to the relatively small number of years in the panel.[5]

An intercept is always included in the regression model (3.1). Prior standard deviations for all of the regression coefficients (β) are set at $\tau_k = 0.5$. This setting provides support for coefficient values different from zero, but does not support unrealistically large values. In addition, the variance-inflation parameter is set at $\kappa = 4$ for all runs involving outlier detection (see (3.2)). This value is chosen so that there is a discernable difference between the distributions for outliers and non-outliers.

Following most of the current Bayesian literature, integrations are performed by means of the Gibbs sampler (Brooks, and Gelfand and Smith). Given regression model (3.1) and the assumed prior distributions, the Gibbs sampler for this problem has six major components: (1) simulation of the main outlier distribution hyperparameter (η), (2) simulation of the main error variance (σ_ε^2), (3) simulation of the annual random effect variance (σ_y^2), (4) simulation of the annual random effects (y_t), (5) simulation of the individual observation outlier detection parameter ($\theta_{i,t}$), and (6) simulation of the parameter vector (β). The present Gibbs sampler was designed to handle these simulations in the order given above (see Appendix for further details).

[5]The panel has only 6 years of data, after accounting for lagged and differenced variables in the model.

Within each of the ten separate estimation runs, the Gibbs sampler simulates four chains of 15,000 iterations each, for a total of 60,000 draws per estimation run. Starting values are chosen systematically for convenience. Doing this does not present a problem as long as the chains are "long enough" to achieve convergence. This is true because, from the properties of Markov chain Monte Carlo methods (of which the Gibbs sampler is a special case), the chains will have a unique stationary distribution identical to the target distribution. The first half of each chain (7,500 iterations) is discarded as a burn-in procedure. Convergence is monitored by Gelman and Rubin's R-statistic, $\sqrt{\hat{R}}$ (Gelman et al., p. 331-332). Parameters are examined to check convergence, including the annual random effects (y_t), the variance components (σ_ε^2 and σ_y^2), and the outlier detection (η) parameters.

Simulation programs are written in C++ and compiled using Borland C++ Builder 3. The distribution subroutines are C++ programs contained in the SUM module of the M++ Version 7.0 libraries from Dyad Software Corporation. A typical run would last five hours on a personal computer with a Pentium 500 MHz chip and 256 megabytes of RAM.

Results and Discussion

Because of the numerous prior scenarios analyzed, attention is focused first on the model with 50 percent prior for variable inclusion and 10 percent for outliers (Var50Out10). Results for this estimation are given in Table 2. For all of the variables selected for the model in at least fifteen percent of the iterations, $\sqrt{\hat{R}}$ is less than 1.15 in the main estimation, so convergence is assumed. The last column of Table 2 reveals that seven of the variables considered are chosen at least fifty percent of the time. These are the change in crop output, the change in the cost of capital, lagged current asset values, lagged off-farm income, and the cross-product terms for q and off-

Table 2. Summary of the results for the Var50Out10 model.

Variable	Posterior Mean	Posterior Quantiles 2.5%	50%	97.5%	$\sqrt{\hat{R}}$	% of Times Chosen
Intercept	0.015	−0.041	0.015	0.069	1.06	100.00
$I_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	3.08
$q_{i,t}$	0.005	0.000	0.000	0.034	1.01	19.03
$\Delta CQ_{i,t}$	0.012	0.000	0.012	0.036	1.01	58.30
$\Delta LQ_{i,t}$	0.000	0.000	0.000	0.000	1.09	3.19
$\Delta C_{i,t}$	−0.054	−0.084	−0.055	0.000	1.01	95.64
$AGE_{i,t-1}$	0.000	0.000	0.000	0.000	1.08	2.38
$CA_{i,t-1}$	0.016	0.010	0.016	0.024	1.03	100.00
$CL_{i,t-1}$	0.000	0.000	0.000	0.000	1.02	1.31
$NW_{i,t-1}$	0.000	−0.003	0.000	0.000	1.07	7.94
$TL_{i,t-1}$	0.000	0.000	0.000	0.000	1.06	0.44
$OFI_{i,t-1}$	0.025	0.000	0.019	0.081	1.00	52.82
$NCF_{i,t-1}$	0.000	0.000	0.000	0.000	1.03	1.70
$\Delta C_{i,t}^2$	0.000	0.000	0.000	0.000	1.03	2.33
$AGE_{i,t-1}^2$	0.000	0.000	0.000	0.000	1.00	0.01
$CA_{i,t-1}^2$	0.000	0.000	0.000	0.000	1.00	0.02
$CL_{i,t-1}^2$	0.000	0.000	0.000	0.000	1.11	0.55
$NW_{I,t-1}^2$	0.000	0.000	0.000	0.000	1.00	0.01
$TL_{i,t-1}^2$	0.000	0.000	0.000	0.000	1.00	0.06
$OFI_{i,t-1}^2$	0.000	0.000	0.000	0.004	1.01	5.63
$NCF_{i,t-1}^2$	0.000	0.000	0.000	0.000	1.00	0.44
$q_{i,t}\times AGE_{i,t-1}$	0.000	0.000	0.000	0.003	2.38	2.16
$q_{i,t}\times CA_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.28
$q_{i,t}\times CL_{i,t-1}$	0.000	0.000	0.000	0.000	1.02	3.42
$q_{i,t}\times NW_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.19
$q_{i,t}\times TL_{i,t-1}$	0.000	0.000	0.000	0.000	2.47	0.68
$q_{i,t}\times OFI_{i,t-1}$	0.075	0.000	0.085	0.144	1.01	80.70
$q_{i,t}\times NCF_{i,t-1}$	0.000	0.000	0.000	0.000	1.19	1.72
$\Delta CQ_{i,t}\times \Delta LQ_{i,t}$	−0.001	−0.020	0.000	0.000	1.14	7.46
$\Delta CQ_{i,t}\times \Delta C_{i,t}$	−0.001	−0.020	0.000	0.000	2.80	2.81
$\Delta CQ_{i,t}\times AGE_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.03
$\Delta CQ_{i,t}\times CA_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.02
$\Delta CQ_{i,t}\times CL_{i,t-1}$	0.006	0.000	0.007	0.010	1.13	58.62
$\Delta CQ_{i,t}\times NW_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.02
$\Delta CQ_{i,t}\times TL_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.05
$\Delta CQ_{i,t}\times OFI_{i,t-1}$	−0.010	−0.039	0.000	0.000	1.01	38.03
$\Delta CQ_{i,t}\times NCF_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.19
$\Delta LQ_{i,t}\times \Delta C_{i,t}$	0.000	0.000	0.000	0.000	1.00	2.09
$\Delta LQ_{i,t}\times AGE_{i,t-1}$	0.000	0.000	0.000	0.000	1.37	2.44

Table 2. (Continued).

Variable	Posterior Mean	Posterior Quantiles 2.5%	50%	97.5%	$\sqrt{\hat{R}}$	% of Times Chosen
$\Delta LQ_{i,t}\times CA_{i,t-1}$	0.000	0.000	0.000	0.000	1.02	1.05
$\Delta LQ_{i,t}\times CL_{i,t-1}$	0.000	0.000	0.000	0.000	1.04	1.28
$\Delta LQ_{i,t}\times NW_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.33
$\Delta LQ_{i,t}\times TL_{i,t-1}$	0.000	0.000	0.000	0.000	1.02	3.51
$\Delta LQ_{i,t}\times OFI_{i,t-1}$	0.008	−0.010	0.000	0.087	1.00	18.96
$\Delta LQ_{i,t}\times NCF_{i,t-1}$	0.000	0.000	0.000	0.000	1.05	2.82
$\Delta C_{i,t}\times AGE_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.29
$\Delta C_{i,t}\times CA_{i,t-1}$	0.000	0.000	0.000	0.000	1.02	1.04
$\Delta C_{i,t}\times CL_{i,t-1}$	0.002	0.000	0.000	0.033	1.43	14.14
$\Delta C_{i,t}\times NW_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.28
$\Delta C_{i,t}\times TL_{i,t-1}$	0.000	0.000	0.000	0.000	1.05	0.69
$\Delta C_{i,t}\times OFI_{i,t-1}$	0.056	0.000	0.058	0.084	1.05	92.36
$\Delta C_{i,t}\times NCF_{i,t-1}$	0.000	−0.008	0.000	0.000	1.08	4.28
$AGE_{i,t-1}\times CA_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.01
$AGE_{i,t-1}\times CL_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.27
$AGE_{i,t-1}\times NW_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.01
$AGE_{i,t-1}\times TL_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.01
$AGE_{i,t-1}\times OFI_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.38
$AGE_{i,t-1}\times NCF_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.05
$CA_{i,t-1}\times CL_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.32
$CA_{i,t-1}\times NW_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.01
$CA_{i,t-1}\times TL_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.01
$CA_{i,t-1}\times OFI_{i,t-1}$	−0.002	-0.026	0.000	0.000	1.08	12.54
$CA_{i,t-1}\times NCF_{I,t-1}$	0.000	0.000	0.000	0.000	1.00	0.06
$CL_{i,t-1}\times NW_{i,t-1}$	0.000	0.000	0.000	0.000	1.26	0.62
$CL_{i,t-1}\times TL_{i,t-1}$	0.000	0.000	0.000	0.000	1.42	0.79
$CL_{i,t-1}\times OFI_{i,t-1}$	−0.001	−0.024	0.000	0.000	1.21	7.70
$CL_{i,t-1}\times NCF_{i,t-1}$	0.003	0.000	0.000	0.014	1.12	35.26
$NW_{i,t-1}\times TL_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.04
$NW_{i,t-1}\times OFI_{i,t-1}$	0.000	−0.005	0.000	0.000	1.01	7.55
$NW_{i,t-1}\times NCF_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.04
$TL_{i,t-1}\times OFI_{i,t-1}$	0.000	0.000	0.000	0.000	1.02	1.12
$TL_{i,t-1}\times NCF_{i,t-1}$	0.000	0.000	0.000	0.000	1.00	0.12
$OFI_{i,t-1}\times NCF_{i,t-1}$	0.001	0.000	0.000	0.017	1.02	7.64
Hyperparameter						
η	0.801	0.769	0.802	0.833	1.03	
σ_ε^2	0.015	0.013	0.015	0.016	1.01	
σ_γ^2	0.009	0.005	0.008	0.015	1.00	

farm income, change in crop output and current liabilities, and change in the cost of capital and off-farm income. Of these, the change in the cost of capital, lagged current asset values, and the cross-product for the change in the cost of capital and off-farm income are chosen over ninety percent of the time. The linear term for the q model is chosen less than twenty percent of the time. The signs for the linear terms of the chosen variables are as expected. Remarkably, 59 of the 72 variables are selected less than ten percent of the time.

The posterior mean estimate for η indicates that roughly twenty percent of the observations are classified as outliers by the model. The variance for the annual random effects is approximately 0.009, very near the prior and the classical estimate for the same measure. This result is expected given the informative prior placed on the random effects variance. The error variance is estimated to be 0.015, nearly twice the size of the random annual effects variance.

To quickly summarize the results of the other estimation runs and to facilitate comparison among them, Table 3 presents a composite summary. Table 3 lists, for each specification, the variables selected at a greater percentage than the prior for that specification, the mean values of their parameters, the mean values of the variance components and outlier detection hyperparameter, and the outlier detection histogram. In addition, variables selected at least 90 percent of the time are shown in bold characters.

Several definite patterns can be seen in Table 3. First, posterior mean parameter estimates for the variance components and the outlier hyperparameter are very similar for all of the Bayesian estimations with both variable selection and nonzero prior probability of outliers. For any given prior probability of variable selection, the estimation procedure selects nearly the same set of variables regardless of the prior chosen for the proportion of outliers (except for the extreme case of a zero prior proportion of outliers) and the posterior mean values for these parameters are consistent.

Only the value of current assets is chosen at least ninety percent of the time in all of the estimations. Parameter estimates for this variable are consistent across all of the estimations with both the variable selection and outlier detection components. When the outlier detection

Table 3. Composite summary table for Bayesian estimation results.

	Prior Specification[a]									
	Var10 Out10	Var10 Out50	Var10 Out90	Var50 Out10	Var50 Out50	Var50 Out90	Var90 Out0	Var90 Out10	Var90 Out50	Var90 Out90
	Posterior Mean Parameter Value[b]									
Variable										
$\Delta CQ_{i,t}$				0.012	0.015	0.014	**0.045**			
$\Delta C_{i,t}$	−0.013	−0.015	−0.018	**−0.054**	**−0.055**	**−0.056**		**−0.049**	**−0.050**	**−0.045**
$CA_{i,t-1}$	**0.015**	**0.015**	**0.015**	**0.016**	**0.017**	**0.017**	**0.045**	**0.018**	**0.018**	**0.019**
$NW_{i,t-1}$		−0.000	−0.000							
$OFI_{i,t-1}$				0.025	0.024	0.024		**0.056**	**0.060**	**0.064**
$q_{i,t}\times OFI_{i,t-1}$	0.021	0.021	0.022	0.075	0.074	0.068				
$\Delta CQ_{i,t}\times CL_{i,t-1}$	0.002	0.002	0.004	0.006						
$\Delta CQ_{i,t}\times OFI_{i,t-1}$							**−0.057**			
$\Delta LQ_{i,t}\times OFI_{i,t-1}$							**0.109**			
$\Delta C_{i,t}\times CL_{i,t-1}$	0.003	0.003								
$\Delta C_{i,t}\times TL_{i,t-1}$							**0.024**			
$\Delta C_{i,t}\times OFI_{i,t-1}$	0.009	0.010	0.014	**0.056**	**0.055**	**0.055**	**0.137**	**0.069**	**0.069**	**0.072**
$\Delta C_{i,t}\times NCF_{i,t-1}$							**−0.078**			
$AGE_{i,t-1}\times CL_{i,t-1}$							**0.004**			
$CA_{i,t-1}\times OFI_{i,t-1}$							**−0.086**			
$CL_{i,t-1}\times TL_{i,t-1}$							**−0.010**			
$CL_{i,t-1}\times NCF_{i,t-1}$	0.009	0.009	0.007			0.007				
$TL_{i,t-1}\times OFI_{i,t-1}$							**0.029**			
Hyperparameter										
σ_ε^2	0.015	0.015	0.014	0.015	0.015	0.014	0.059	0.014	0.014	0.013
σ_y^2	0.009	0.009	0.009	0.009	0.009	0.009	0.009	0.009	0.009	0.009
η	0.808	0.803	0.780	0.801	0.797	0.774		0.794	0.789	0.763
Outlier %[c]	Number of Observations									
0	0	0	0	0	0	0	2196	0	0	0
(0, 10]	1482	1448	1299	1439	1409	1256	0	1372	1338	1159
(10, 20]	302	327	426	318	335	446	0	353	380	502
(20, 80]	210	216	249	230	242	263	0	260	263	303
(80, 90]	42	40	36	32	28	37	0	29	29	33
(90, 100]	160	165	186	177	182	194	0	182	186	199

[a]For the Prior Specifications, "VarXOutY" means that results correspond to an X percent prior probability that the variables are included in the model, and a prior that Y percent of the observations are outliers. Thus, for example, Var10Out50 means a 10 percent prior probability that the variables are included in the model, and a prior that 50 percent of the observations are outliers.

[b]If there is no value in a cell, either the variable was not selected X percent (the prior percentage) of the time or the parameter was not estimated in that scenario. Numbers in **bold** indicate that the variable was selected at least 90 percent of the time.

[c]The percentage of times the observation was chosen as an outlier. For example, for the Var50Out10 scenario, 1439 (65.5 percent) of the observations were not selected as outliers over 90 percent of the iterations; 1864 (84.9 percent) of the observations were not selected as outliers over 70 percent of the time, but 177 (8.1 percent) of the observations were selected as outliers over 90 percent of the time.

component is removed (Out0), the estimate more than doubles in size. The cross-product term for the change in the cost of capital and off-farm income is chosen at least ninety percent of the time for all estimation where the prior probability of variable inclusion is at or above fifty percent (Var50 and Var90). For the specifications with both variable selection and outlier detection components, the parameter estimates for this cross-product are similar, with each being contained in the quantile intervals (2.5 to 97.5 percent) from the other estimations. The parameter estimate for this cross-product in the no outlier specification is roughly double that from the other specifications. The change in the cost of capital also appears in all of the specifications with both components, but is not chosen in the no outlier specification. If the prior probability of variable inclusion is at or above fifty percent, the change in the cost of capital is chosen over ninety percent of the time and again, the parameter estimates are similar across these specifications.

The outlier histograms reported in the bottom half of Table 3 are also extremely robust across nonzero priors. In the scenarios with nonzero prior probability of outliers, between 1,661 and 1,784 observations (75.6 and 81.2 percent of the total observations, respectively) were not selected as outliers in over 80 percent of the iterations, and between 160 and 199 observations (7.3 and 9.1 percent of the total observations, respectively) were selected as outliers over 90 percent of the iterations.

For specifications with both the variable selection and outlier detection components, those variables chosen at least ninety percent of the time have very similar parameter estimates. To examine the effects of the addition of the outlier detection component, the no outlier specification (Out0) is compared to the other models with the 90 percent prior probability for variable inclusion (Var90Out10, Var90Out50, and Var90Out90). First, the no outlier specification included eleven more variables than and excluded two of the four variables chosen in the other 90 percent prior variable inclusion specifications. Second, the parameter estimates in the no outlier case for the two variables in common across all four specifications are at least double the size of the estimates from the others. Third, the posterior mean of the error variance is four times greater under the no outlier case versus the scenarios allowing for outliers. This is to be expected, because the error variance must be large enough to accommodate all the outliers that are not classified as such. Overall, it can be concluded that the choice of nonzero priors for the variable selection and outlier detection components has a much smaller effect on the results from the Bayesian analysis than the addition of an outlier detection component.

Despite the aforementioned differences, the no outlier and outlier detection scenarios held a couple of similarities. The value of current assets and the cross-product of the cost of capital and off-farm income strongly appear in all four specifications. The q model is not represented in any of the specifications. Also, the posterior means of the random effects variance are quite similar.

Comparison with Results from Classical Analysis

At this point, it is of interest to compare the results from the Bayesian approach with those obtained under classical methods. To this end, (3.1) was fitted using maximum likelihood under the mixed model procedure in SAS 6.12 for Windows. Parameter estimates are reported in Table 4. Approximately 40 percent (30 of 73) of the parameter estimates are significantly different from zero at the 5 percent level. Five of the 11 linear effects are significant and all of those are positively related to investment. For the eight squared terms, only the change in the cost of capital and current liabilities are significantly different from zero at the five percent level. Of the 52 cross effects, 21 have parameter estimates significantly different from zero, including all but two ($q_{i,t} \times CA_{i,t-1}$ and $\Delta LQ_{i,t} \times CA_{i,t-1}$) of the cross-products involving current assets. The variance estimates indicate that the residual error dominates the annual random effect.

It is clear from Table 4 that the q model of investment fares quite poorly under the classical regression framework. Only one ($q_{i,t} \times AGE_{i,t-1}$) of the eight terms involving $q_{i,t}$ is different from zero at the 5 percent level of significance. This poor performance of the q model is not the result of $q_{i,t}$ being unable to explain investment by itself. Rather, it is due to the fact that $q_{i,t}$ adds little to explain investment, *once other explanatory variables are taken into*

Table 4. Classical regression results.

Variable	Estimate	Std. Error	Variable	Estimate	Std. Error
Intercept	0.090*	0.016	$\Delta LQ_{i,t} \times CA_{i,t-1}$	−0.0035	0.0082
$I_{i,t-1}$	0.001	0.017	$\Delta LQ_{i,t} \times CL_{i,t-1}$	0.006	0.012
$q_{i,t}$	0.044	0.024	$\Delta LQ_{i,t} \times NW_{i,t-1}$	0.0073*	0.0033
$\Delta CQ_{i,t}$	0.035**	0.013	$\Delta LQ_{i,t} \times TL_{i,t-1}$	0.0063	0.0058
$\Delta LQ_{i,t}$	0.025*	0.011	$\Delta LQ_{i,t} \times OFI_{i,t-1}$	0.150**	0.046
$\Delta C_{i,t}$	−0.057	0.030	$\Delta LQ_{i,t} \times NCF_{i,t-1}$	0.006	0.016
$AGE_{i,t-1}$	0.0041*	0.0020	$\Delta C_{i,t} \times AGE_{i,t-1}$	0.0014	0.0017
$CA_{i,t-1}$	0.0329**	0.0083	$\Delta C_{i,t} \times CA_{i,t-1}$	0.0221*	0.0098
$CL_{i,t-1}$	0.030*	0.013	$\Delta C_{i,t} \times CL_{i,t-1}$	−0.005	0.022
$NW_{i,t-1}$	−0.0010	0.0024	$\Delta C_{i,t} \times NW_{i,t-1}$	0.0024	0.0027
$TL_{i,t-1}$	−0.0110	0.0060	$\Delta C_{i,t} \times TL_{i,t-1}$	0.0114	0.0091
$OFI_{i,t-1}$	0.081*	0.037	$\Delta C_{i,t} \times OFI_{i,t-1}$	0.093**	0.029
$NCF_{i,t-1}$	0.028	0.016	$\Delta C_{i,t} \times NCF_{i,t-1}$	−0.085**	0.021
$\Delta C_{i,t}^2$	−0.034*	0.017	$AGE_{i,t-1} \times CA_{i,t-1}$	0.0028*	0.0011
$AGE_{i,t-1}^2$	−0.00018	0.00010	$AGE_{i,t-1} \times CL_{i,t-1}$	−0.0000	0.0020
$CA_{i,t-1}^2$	0.0046	0.0030	$AGE_{i,t-1} \times NW_{i,t-1}$	0.00040	0.00027
$CL_{i,t-1}^2$	−0.0226**	0.0082	$AGE_{i,t-1} \times TL_{i,t-1}$	0.00097	0.00088
$NW_{i,t-1}^2$	0.00026	0.00024	$AGE_{i,t-1} \times OFI_{i,t-1}$	0.0013	0.0035
$TL_{i,t-1}^2$	0.0013	0.0016	$AGE_{i,t-1} \times NCF_{i,t-1}$	−0.0061**	0.0022
$OFI_{i,t-1}^2$	0.002	0.035	$CA_{i,t-1} \times CL_{i,t-1}$	0.0196*	0.0080
$NCF_{i,t-1}^2$	0.0127	0.0085	$CA_{i,t-1} \times NW_{i,t-1}$	−0.0050**	0.0018
$q_{i,t} \times AGE_{i,t-1}$	0.0083**	0.0031	$CA_{i,t-1} \times TL_{i,t-1}$	−0.0112**	0.0033
$q_{i,t} \times CA_{i,t-1}$	−0.0167	0.0098	$CA_{i,t-1} \times OFI_{i,t-1}$	−0.105**	0.026
$q_{i,t} \times CL_{i,t-1}$	−0.022	0.020	$CA_{i,t-1} \times NCF_{i,t-1}$	0.0250**	0.0085
$q_{i,t} \times NW_{i,t-1}$	−0.0001	0.0048	$CL_{i,t-1} \times NW_{i,t-1}$	−0.0063*	0.0030
$q_{i,t} \times TL_{i,t-1}$	0.0102	0.0096	$CL_{i,t-1} \times TL_{i,t-1}$	−0.0017	0.0057
$q_{i,t} \times OFI_{i,t-1}$	−0.021	0.084	$CL_{i,t-1} \times OFI_{i,t-1}$	−0.031	0.047
$q_{i,t} \times NCF_{i,t-1}$	−0.050	0.028	$CL_{i,t-1} \times NCF_{i,t-1}$	0.008	0.014
$\Delta CQ_{i,t} \times \Delta LQ_{i,t}$	−0.022	0.016	$NW_{i,t-1} \times TL_{i,t-1}$	0.0022*	0.0011
$\Delta CQ_{i,t} \times \Delta C_{i,t}$	0.007	0.013	$NW_{i,t-1} \times OFI_{i,t-1}$	−0.0063	0.0055
$\Delta CQ_{i,t} \times AGE_{i,t-1}$	0.0013	0.0013	$NW_{i,t-1} \times NCF_{i,t-1}$	−0.0036	0.0028
$\Delta CQ_{i,t} \times CA_{i,t-1}$	−0.0121**	0.0042	$TL_{i,t-1} \times OFI_{i,t-1}$	0.027*	0.012
$\Delta CQ_{i,t} \times CL_{i,t-1}$	−0.0022	0.0086	$TL_{i,t-1} \times NCF_{i,t-1}$	−0.0097	0.0060
$\Delta CQ_{i,t} \times NW_{i,t-1}$	−0.0002	0.0018	$OFI_{i,t-1} \times NCF_{i,t-1}$	0.103*	0.053
$\Delta CQ_{i,t} \times TL_{i,t-1}$	0.0038	0.0041			
$\Delta CQ_{i,t} \times OFI_{i,t-1}$	−0.085**	0.025	Variance Parameters		
$\Delta CQ_{i,t} \times NCF_{i,t-1}$	0.033**	0.012	$\hat{\sigma}_{\varepsilon}^2$	0.057	
$\Delta LQ_{i,t} \times \Delta C_{i,t}$	0.014	0.016	$\hat{\sigma}_{y}^2$	0.001	
$\Delta LQ_{i,t} \times AGE_{i,t-1}$	−0.0069**	0.0020	R^2	0.302	

* (**) Significant at the 5 (1) percent level, based on the two-sided *t*-test.

account. This is true because $q_{i,t}$ is highly significant in the following univariate regression, analogous to that reported in Table 4 (standard errors within parentheses):

$$(4.1)\quad \underset{}{I_{i,t}/K_{i,t-1}} = \underset{(0.017)}{0.070} + \underset{(0.012)}{0.103}\, q_{i,t}, \qquad \hat{\sigma}_\varepsilon^2 = 0.075,\ \hat{\sigma}_y^2 = 0.001,\ R^2 = 0.108.$$

The Bayesian specification the most similar to the classical analysis is the Var90Out0 specification. In comparing these results, it can be seen that the classical analysis found many more variables to be significant. Of the eleven variables chosen in the Var90Out0 specification, eight were also found to be significant in the classical analysis. The signs of the parameters for these eight variables are the same across the analyses, but the magnitudes differ. Also, the two variables that are chosen by all four Bayesian specifications with a ninety prior probability for variable inclusion, lagged current assets and the cross-product for the change in the cost of capital and off-farm income, are found to be significant in the classical analysis.

Expected Change Estimates

From an economic standpoint, it is of interest to examine the absolute impacts of the factors on farm machinery investment. To this end, Table 5 reports the estimated expected changes in the investment rate related to a one unit change in each regressor from the classical results, and the marginal posterior distributions of the expected change in the investment rate from the Bayesian results. Because of the skewness in the investment data (see Table 1), expected changes were computed at both the mean and median values for all variables.

The Bayesian approach taken in this manuscript allows us to approximate the marginal posterior distributions of unit changes, because the latter portions of the chains can be thought of as coming from the posterior distributions of interest. Within the Bayesian framework, we then obtain point estimates and credible intervals (the Bayesian equivalent to frequentist confidence intervals) for the expected changes. This approach also allows us to incorporate the uncertainty about all model parameters directly into the expected change estimates. From the parameter

Table 5. Summary of expected changes in the investment rate.[a]

Variable	Classical Results Mean Estimate	Bayesian Results Posterior Mean	Posterior Quantiles 2.5%	50%	97.5%	$\sqrt{\hat{R}}$	% of Times Non-Zero
At the mean:							
$I_{i,t-1}$	0.001	0.000	0.000	0.000	0.000	1.00	3.08
$q_{i,t}$	0.044	0.005	0.000	0.000	0.034	1.01	19.03
$\Delta CQ_{i,t}$	0.035	0.012	0.000	0.012	0.036	1.01	58.30
$\Delta LQ_{i,t}$	0.025	0.000	0.000	0.000	0.000	1.09	3.19
$\Delta C_{i,t}$	−0.057	−0.054	−0.084	−0.055	0.000	1.01	95.64
$AGE_{i,t-1}$	0.004	0.000	0.000	0.000	0.000	1.08	2.38
$CA_{i,t-1}$	0.033	0.016	0.010	0.016	0.024	1.03	100.00
$CL_{i,t-1}$	0.030	0.000	0.000	0.000	0.000	1.02	1.31
$NW_{i,t-1}$	−0.001	0.000	−0.003	0.000	0.000	1.07	7.94
$TL_{i,t-1}$	−0.011	0.000	0.000	0.000	0.000	1.06	0.44
$OFI_{i,t-1}$	0.081	0.025	0.000	0.019	0.081	1.00	52.82
$NCF_{i,t-1}$	0.028	0.000	0.000	0.000	0.000	1.03	1.70
At the median:							
$I_{i,t-1}$	0.001	0.000	0.000	0.000	0.000	1.00	3.08
$q_{i,t}$	0.042	−0.004	−0.015	−0.008	0.028	1.00	86.08
$\Delta CQ_{i,t}$	0.041	0.011	−0.003	0.011	0.037	1.01	85.48
$\Delta LQ_{i,t}$	0.010	−0.001	−0.009	0.000	0.005	1.02	33.11
$\Delta C_{i,t}$	-0.076	−0.060	−0.091	−0.061	−0.011	1.01	99.30
$AGE_{i,t-1}$	0.003	0.000	0.000	0.000	0.000	1.08	7.33
$CA_{i,t-1}$	0.037	0.016	0.010	0.016	0.024	1.04	100.00
$CL_{i,t-1}$	0.047	−0.001	−0.003	0.000	0.002	1.06	95.96
$NW_{i,t-1}$	0.002	0.000	−0.003	0.000	0.001	1.07	16.30
$TL_{i,t-1}$	-0.014	0.000	0.000	0.000	0.000	1.04	6.77
$OFI_{i,t-1}$	0.116	0.022	−0.007	0.015	0.084	1.00	99.35
$NCF_{i,t-1}$	0.027	−0.001	−0.005	0.000	0.001	1.06	45.57

[a]The Bayesian results are based on the Var50Out10 specification.

draws from the Gibbs sampler for the Var50Out10 case (which had prior probabilities of 50 percent for the variable inclusion component and of 10 percent for the outlier detection component), marginal posterior distributions of the expected changes were computed.

The latter columns of Table 5 summarize the Bayesian results (i.e., the posterior means, selected posterior quantiles, Gelman and Rubin's R-statistics, and the percentage of times the estimate is non-zero out of 60,000 iterations). At the mean values, the change in the cost of capital has the largest impact followed by off-farm income, the value of current assets, the change in crop output, and q. The other variables have a negligible impact. Non-zero estimates for these expected changes only occur less than ten percent of the time. At the median values, the order is the same but the strength of the impacts varies. The impact from q switches from positive to negative. Also, quite small negative impacts are seen from the change in livestock output, lagged current liabilities, and net cash flow.

The classical point estimates in Table 4 were used to calculate the expected change in the investment rate related to a one unit change in each regressor. These are shown in the second column in Table 5. The median value of current investment is negative, indicating real disinvestment on the farm. For both mean and median values, lagged off-farm income has the largest impact on the expected value of the investment rate. The positive impact indicates that as the off-farm income rises, the farm machinery investment rate also rises. The change in the cost of capital has the next largest impact. For mean values, it is followed by q, change in crop output, and lagged current assets. For median values, it is followed by lagged current liabilities, q, and change in crop output.

When comparing these figures from the two approaches, several important differences can be observed. The classical estimates indicate stronger impacts from almost all of the variables. The Bayesian estimates show the change in the cost of capital to have the largest impacts, whereas the classical estimates point to off-farm income. Lagged current assets are near the middle of the pack in the classical results, but are the third highest in the Bayesian approach.

Conclusions

It is widely agreed that private business investment plays a crucial role in enhancing economic performance/growth. It is a major contributor to long-term growth and short-term fluctuations in economic activity. Many studies have examined, both theoretically and empirically, private investment. Two of the major models from this literature are the q and neoclassical models. However, many studies have found that firms' internal financial variables also have a significant impact on investment.

Our analysis examines three stylized facts that emerge from the literature. First, *there are no empirical studies testing whether the q model performs better than the neoclassical paradigm, or vice versa, in explaining agricultural investment*. Second, *the specific financial variables used to explain investment vary substantially across studies*. Third, *none of the existing studies accounts for outlier effects*.

The present study employs a Bayesian approach to address these issues. Specifically, we form an investment rate composite model including factors from the q and neoclassical models and financial variables. The Bayesian approach is constructed with model selection and outlier detection components. This contribution transcends investment research in that the advocated techniques can be applied to the empirical analysis of many other economic issues. The techniques are based on very recent works in Bayesian model selection, and the present study is one of the first to employ them in an econometric setting.

For the panel data analyzed, consisting of 366 Iowa farms over the period 1991-1998, more support is found for the neoclassical model than for the q model of investment. The change in the cost of capital is selected for the investment model both in a linear form and within a cross-product with off-farm income. Confirming the findings of the previous literature, a farm's financial situation is found to affect its investment significantly. This is true because lagged current asset values and off-farm income add greatly to the investment model. In fact, lagged current asset value is the only variable that is chosen at least ninety percent of the time by all of the Bayesian specifications and is also found to be significant at the 95 percent level in the classical analysis. Other standard financial variables, such as net worth and total liabilities, are virtually ignored by both estimation techniques.

The incorporation of an outlier detection component changes the results drastically, in both the variables chosen and the magnitudes of the estimated parameter values. In the present sample, it was found that roughly twenty percent of the observations are classified as outliers. The classical analysis of the model (not accounting for outliers) found 29 variables to be statistically significant at the 95 percent level. The Bayesian analysis with the model selection, but not the outlier detection, component chose only 11 variables for the model, including eight variables that were significant in the classical analysis. The Bayesian analyses with both components indicated that only four variables, the linear terms for the change in the cost of capital, lagged current assets, and lagged off-farm income and the cross-product for the change in the cost of capital and off-farm income, are important in the investment rate model. The shifts in the variable selection are significant in that this also implies that the inferences drawn change.

To inspect this more closely, expected changes in investment rates due to changes in the explanatory variables were also examined. It was found that the Bayesian and classical estimates differ dramatically. In almost all cases, the classical estimates exceed the Bayesian estimates. Also, the rankings of the effects of the explanatory variables shift between the analyses. These results and the nature of most investment data suggest that past investment work that does not account for outliers may have led to inaccurate inferences.

Appendix A: Output and Output Price Indices

The formulas for the output measures and the output price index are given in (A1) and (A2), respectively:

(A1) $O_{i,t} = \Sigma_j \overline{p}_j \ A_{j,i,t}$,

(A2) $$p_t^O = \frac{\Sigma_j p_{j,t} \ \overline{A}_j}{\Sigma_j p_{j,1998} \ \overline{A}_j},$$

where $\overline{p}_j$ is the 1991-98 average price for commodity j, $A_{j,i,t}$ is the quantity of commodity j produced by farm i in year t, $p_{j,t}$ is the price of commodity j in year t, and $\overline{A}_j$ is the 1991-98 average farm-level quantity produced of commodity j. All prices employed in these calculations are deflated by the "Producer Price Index for Farm Products."

Appendix B: Posterior Conditional Distributions

Given the assumed prior distributions for η, σ_ε^2, σ_y^2, y_t, and $\theta_{i,t}$, their posterior conditional distributions are as shown in (B.1), (B.2), (B.3), (B.4), and (B.5), respectively:

(B.1) $\eta \mid I_{i,t}/K_{i,t-1}, y_t, \beta, X_{i,t}, \sigma_\varepsilon^2, \sigma_y^2, \theta_{i,t} \sim \text{Beta}\left(\Sigma_{i=1}^n \Sigma_{t=1}^T \theta_{i,t} + \gamma, nT - \Sigma_{i=1}^n \Sigma_{t=1}^T \theta_{i,t} + \varphi\right)$,

(B.2) $$\sigma_\varepsilon^2 \mid I_{i,t}/K_{i,t-1}, y_t, \beta, X_{i,t}, \eta, \sigma_y^2, \theta_{i,t} \sim \text{Inverse-Gamma}\left(0.5\, n\, T,\ 0.5\, \Sigma_{i=1}^n \Sigma_{t=1}^T \frac{\left(I_{i,t}/K_{i,t-1} - \beta X_{i,t} - y_t\right)^2}{\left(\theta_{i,t} + \kappa^2 - \kappa^2 \theta_{i,t}\right)}\right),$$

(B.3) $\sigma_y^2 \mid I_{i,t}/K_{i,t-1}, y_t, \beta, X_{i,t}, \eta, \sigma_\varepsilon^2, \theta_{i,t} \sim \text{Inverse-Gamma}\left(0.5(T + n_0), 0.5\left(n_0 \sigma_0^2 + \Sigma_{t=1}^T y_t^2\right)\right)$,

(B.4) $$y_t \mid I_{i,t}/K_{i,t-1}, \beta, X_{i,t}, \eta, \sigma_\varepsilon^2, \sigma_y^2, \theta_{i,t} \sim \text{Normal}\left(\left(\frac{\sigma_y^2}{\sigma_\varepsilon^2 + \sigma_y^2 W}\right) \Sigma_{i=1}^n \frac{\left(I_{i,t}/K_{i,t-1} - \beta X_{i,t}\right)}{\left(\theta_{i,t} + \kappa^2 - \kappa^2 \theta_{i,t}\right)}, \left(\frac{1}{\sigma_\varepsilon^2} W + \frac{1}{\sigma_y^2}\right)^{-1}\right),$$

(B.5) $p(\theta_{i,t} = 0 \mid I_{i,t}/K_{i,t-1}, y_t, \beta, X_{i,t}, \eta, \sigma_\varepsilon^2, \sigma_y^2) = p_0/(p_0 + p_1)$, $p(\theta_{i,t} = 1 \mid \cdot) = p_1/(p_0 + p_1)$, where n denotes number of farms, T is the number of years, $W \equiv \sum_{i=1}^{n} 1/(\theta_{i,t} + \varkappa^2 - \varkappa^2\theta_{i,t})$, $p_0 \equiv \exp\{-0.5\,[(I_{i,t}/K_{i,t-1} - \beta X_{i,t} - y_t)/(\kappa\,\sigma_\varepsilon)]^2\}$, and $p_1 \equiv \kappa\eta/(1-\eta)\exp\{-0.5[(I_{i,t}/K_{i,t-1} - \beta X_{i,t} - y_t)/\sigma_\varepsilon]^2\}$.

The posterior distribution of β_k conditional on $\theta_{i,t}$, $I_{i,t}/K_{i,t-1}$, y_t, $X_{i,t}$, η, σ_ε^2, σ_y^2, and $\beta_{j\neq k}$ originates from the simplified model $z_{i,t} = \beta_k X_{i,t,k} + \varepsilon_{i,t}$, where $z_{i,t} \equiv I_{i,t}/K_{i,t-1} - \sum_{j\neq k}\beta_j X_{i,t,j} - y_t$, and $\varepsilon_{i,t}$ is distributed as shown in (3.2). Given the prior probability that $\beta_k = 0$ ($\underline{p}_k$), the posterior probability that $\beta_k = 0$ is given by $\bar{p}_k = \underline{p}_k/[\underline{p}_k + (1-\underline{p}_k)\,BF_k]$, where BF_k denotes the conditional Bayes factor in favor of $\beta_k \neq 0$ versus $\beta_k = 0$:

$$(B.6)\quad BF_k \equiv \left(\frac{1}{\tau_k}\left(\frac{1}{\sigma_\varepsilon^2}\sum_{i=1}^{n}\sum_{t=1}^{T}\frac{X_{i,t,k}^{2}}{\left(\theta_{i,t}+\varkappa^2-\varkappa^2\theta_{i,t}\right)}+\frac{1}{\tau_k^2}\right)^{-0.5}\right)$$
$$\exp\left(0.5\left(\frac{1}{\sigma_\varepsilon^2}\sum_{i=1}^{n}\sum_{t=1}^{T}\frac{X_{i,t,k}z_{i,t}}{\left(\theta_{i,t}+\varkappa^2-\varkappa^2\theta_{i,t}\right)}\right)^2\left(\frac{1}{\sigma_\varepsilon^2}\sum_{i=1}^{n}\sum_{t=1}^{T}\frac{X_{i,t,k}^{2}}{\left(\theta_{i,t}+\varkappa^2-\varkappa^2\theta_{i,t}\right)}+\frac{1}{\tau_k^2}\right)^{-1}\right).$$

Larger BF_k leads to greater chances for β_k to be included in the regression model. The posterior probability that $\beta_k \neq 0$ is equal to $1-\bar{p}_k$.

Finally, for $\beta_k \neq 0$, the posterior conditional distribution is:

$$(B.7)\quad \beta_k \mid I_{i,t}/K_{i,t-1}, y_t, \beta_{j\neq k}, X_{i,t}, \eta, \sigma_\varepsilon^2, \sigma_y^2, \theta_{i,t} \sim$$

$$\text{Normal}\left(\begin{array}{c}\left(\frac{1}{\sigma_\varepsilon^2}\sum_{i=1}^{n}\sum_{t=1}^{T}\frac{X_{i,t,k}z_{i,t}}{\theta_{i,t}+\kappa^2-\kappa^2\theta_{i,t}}\right)\left(\frac{1}{\sigma_\varepsilon^2}\sum_{i=1}^{n}\sum_{t=1}^{T}\frac{X_{i,t,k}^{2}}{\theta_{i,t}+\kappa^2-\kappa^2\theta_{i,t}}+\frac{1}{\tau_k^2}\right)^{-1}, \\ \left(\frac{1}{\sigma_\varepsilon^2}\sum_{i=1}^{n}\sum_{t=1}^{T}\frac{X_{i,t,k}^{2}}{\theta_{i,t}+\kappa^2-\kappa^2\theta_{i,t}}+\frac{1}{\tau_k^2}\right)^{-1}\end{array}\right).$$

References

Abel, A. B. and O. Blanchard. 1986. "The Present Value of Profits and Cyclical Movements in Investment." *Econometrica* 54:249-273.

Benjamin, C., and E. Phimister. 1997. "Transaction Costs, Farm Finance and Investment." *European Review of Agricultural Economics* 24:453-466.

Bierlen, R. W. and A. M. Featherstone. 1996. "Credit Rationing and Farm Machinery Investment." Manuscript, University of Arkansas, Fayetteville.

Bierlen, R., B. L. Ahrendsen, and B. L. Dixon. 1998. "Impacts of Financial Characteristics and the Boom-Bust Cycle on the Farm Inventory-Cash Flow Relationship." *Journal of Agricultural and Applied Economics* 30:363-377.

Bierlen, R., P. J. Barry, B. L. Dixon, and B. L. Ahrendsen. 1998. "Credit Constraints, Farm Characteristics, and the Farm Economy: Differential Impacts on Feeder Cattle and Beef Cow Inventories." *American Journal of Agricultural Economics* 80:708-723.

Brainard, W. C. and J. Tobin. 1968. "Pitfalls in Financial Model Building." *American Economic Review* 58: 99-122.

Brooks, S. P. 1998. "Markov Chain Monte Carlo Method and its Application." *The Statistician* 47: 69-100.

Chirinko, R. S. 1993. "Business Fixed Investment Spending: A Critical Survey of Modeling Strategies, Empirical Results, and Policy Implications." Working paper RWP 93-01. Federal Reserve Bank of Kansas City.

Gelfand, A. E. and A. F. M. Smith. 1990. "Sampling-Based Approaches to Calculating Marginal Densities." *Journal of the American Statistical Association* 85: 398-409.

Gelman, A., J. B. Carlin, H. S. Stern, and D. B. Rubin. 1995. *Bayesian Data Analysis*. New York: Chapman and Hall.

Geweke, J. K. 1996. "Variable Selection and Model Comparison in Regression." In *Bayesian Statistics 5*, edited by J. M. Bernardo, J. O. Berger, A. P. Dawid, and A. F. M. Smith, pages 609-620. Oxford: Oxford Press.

Gilchrist, S. and C. Himmelberg. 1998. "Investment, Fundamentals and Finance." National Bureau of Economic Research, Working Paper No. 6652.

Herendeen, J. B., and W. Grisley. 1988. "A Dynamic 'q' Model of Investment, Financing and Asset Pricing: An Empirical Test for the Agricultural Sector." *Southern Economic Journal* 55:360-373.

Hubbard, R. G. 1998. "Capital Market Imperfections and Investment." *Journal of Economic Literature* 36:193-225.

Hubbard, R. G. and A. K. Kashyap. 1992. "Internal Net Worth and the Investment Process: An Application to U.S. Agriculture." *Journal of Political Economy* 100: 506-534.

Jensen, F. E., J. S. Lawson, and L. N. Langemeier. 1993. "Agricultural Investment and Internal Cash Flow Variables." *Review of Agricultural Economics* 15: 295-306.

Jorgenson, D. W. 1963. "Capital Theory and Investment Behavior." *American Economic Review* 53: 247-259.

Jorgenson, D. W. 1971. "Econometric Studies of Investment Behavior: A Survey." *Journal of Economic Literature* 9: 1111-1147.

Keynes, J. M. 1936. *The General Theory of Employment, Interest, and Money*. New York: Harcourt Brace.

Knopf, E. and R. Schoney. 1993. "An Evaluation of Farm Financial Benchmarks and Loan Success/Failure: The Case of the Agricultural Credit Corporation of Saskatchewan." *Canadian Journal of Agricultural Economics* 41: 61-69.

Miller, L. H. and E. L. LaDue. 1989. "Credit Assessment Models for Farm Borrowers: A Logit Analysis." *Agricultural Finance Review* 49: 22-36.

Tobin, J. 1969. "A General Equilibrium Approach to Monetary Theory." *Journal of Money, Credit, and Banking* 1: 15-29.

U.S. Department of Agriculture, National Agricultural Statistics Service. *Agricultural Statistics 1999*. Washington, D.C.

U.S. Department of Commerce, Economics and Statistics Administration, Bureau of Economic Analysis. *Survey of Current Business*. Washington, D.C.

U.S. Department of Commerce, Economics and Statistics Administration, Bureau of the Census. *1997 Census of Agriculture*. Washington, D.C.

Walraven, N. A. and M. Post. 1999. *Agricultural Finance Databook*. Division of Research and Statistics, Board of Governors of the Federal Reserve System. Washington, D.C.

Weersink, A. J. and L. W. Tauer. 1989. "Comparative Analysis of Investment Models for New York Dairy Farms." *American Journal of Agricultural Economics* 71: 136-146.